



TITLE:

Development and validation of open-source software for DNA mixture interpretation based on a quantitative continuous model(Abstract_要旨)

AUTHOR(S):

Manabe, Sho

CITATION:

Manabe, Sho. Development and validation of open-source software for DNA mixture interpretation based on a quantitative continuous model. 京都大学, 2018, 博士(医科学)

ISSUE DATE:

2018-03-26

URL:

<https://doi.org/10.14989/doctor.k21024>

RIGHT:

京都大学	博士（医 科 学）	氏 名	眞 鍋 翔
論文題目	Development and validation of open-source software for DNA mixture interpretation based on a quantitative continuous model（定量的連続性モデルに基づく DNA 混合試料解析用オープンソースソフトウェアの開発と検証）		
（論文内容の要旨）			
<p>法医学領域への DNA 分析技術の導入は、身元確認や犯罪捜査に大きな変革をもたらした。現在のわが国の法医学領域で用いられている DNA 検査法は、ヒトゲノム中のマイクロサテライト（short tandem repeat（STR））15 座位を PCR で増幅後、反復単位の繰り返し数をキャピラリー電気泳動システムで判定して型を決定するものである。単一個人由来の DNA であれば、1 座位あたり 1 本（ホモ接合体）または 2 本（ヘテロ接合体）のピークが検出され、型判定は容易である。ただし、法医学領域で扱う試料は、複数のヒト DNA が混合した試料（混合試料）や量的に極めて少ない試料が多い。このような試料では 3 本以上のピークが検出されたり、本来検出されるはずのピークが検出されない場合がある。特に犯罪捜査などでの DNA 鑑定では、「現場試料に何人の DNA が含まれているか」、「被疑者の DNA は現場試料中に含まれているか」を確率を用いて評価する必要があるが、このような試料では困難を極める。</p> <p>混合試料に対する確率的評価について、欧米諸国では様々な方法が提唱され、ソフトウェアも開発されている。中でも quantitative continuous model は、試料に含まれている DNA 量を反映するピークの高さの値を利用するので、厳密な確率的評価を行えるモデルとして注目されている。しかし、現在のソフトウェアの多くは購入に数百万円を要し、プログラムの中身が非公開で不明点も多い。また、計算に使用するパラメータの値が限定的で、検査キットや座位毎の細かな違いまで十分に補正できていない。そこで本研究では、これらの問題点を解決するため、quantitative continuous model に基づく独自の混合試料解析用のソフトウェア（<i>Kongoh</i>）を開発した。また、模擬の混合試料を型判定して、開発したソフトウェアで解析し、その有効性を検証した。</p> <p><i>Kongoh</i> の開発には、プログラミング言語 R（ver. 3.4.1）を用いたが、R に不慣れな人でも操作できるように graphical user interface（GUI）化した。確率的評価のうち「現場試料に何人の DNA が含まれているか」については、<i>Kongoh</i> が自動的に最適な人数（1－4 人）を決定するように工夫した。また、「被疑者の DNA は現場試料中に含まれているか」については、被疑者の DNA が含まれている場合と含まれていない場合に混合試料の型となる尤度の比（尤度比）を算出して、その値が大きいほど含まれていることを強く支持するものとした。これらの計算を厳密に行うために、ピーク高に影響を及ぼすパラメータを 5 種類設定し、多数の実験データ（計 234 サンプル）に基づいて座位毎の細かな違いを推定した上で <i>Kongoh</i> に組み入れた。</p> <p><i>Kongoh</i> の検証は、世界的に推奨されている米国のガイドラインに基づいて行った。検証に用いた試料は 2 人から 4 人までの混合試料（計 80 サンプル）であり、全 DNA 量や各個人の DNA の混合比、DNA 劣化の程度について様々なパターンを含むようにした。その結果、不検出ピークが 5 個未満であれば、試料に</p>			

DNA が含まれている個人の尤度比は概ね 10,000 以上となり、含まれていることを支持した。反対に DNA が含まれていない個人について、各試料につき 100 名ずつ尤度比を算出すると全て 1 未満となり、含まれていないことを支持した。また、「現場試料に何人の DNA が含まれているか」を推定すると、その正答率は 2 人の混合試料で 92.5%、3 人の混合試料で 81.5%と高い正答率を示した。また、他の quantitative continuous model のソフトウェアと計算結果を比較したところ、「現場試料に何人の DNA が含まれているか」について *Kongoh* の方が高い正答率を示した。

以上より、ソフトウェア *Kongoh* は混合試料の確率的評価に有用であることが示された。*Kongoh* は誰でも無料で使用することができ、わが国の検査システムに即したソフトウェアであるため、法医実務での普及が期待される。

（論文審査の結果の要旨）

法医学領域における最近の DNA 鑑定では、事件の被疑者等の問題となる人の DNA が複数人の DNA が混合した試料に存在するか否かを評価する必要が急増している。しかし、わが国ではその評価が殆ど行われていない。欧米諸国では混合試料を解析するためのソフトウェアが既開発されているが、多くは購入に数百万円を要し、プログラムに不明な箇所がある点やあいまいなパラメータ値を計算に使用している点が問題である。そこで本研究では、これらの問題点を解決すべく、混合試料を解析するソフトウェアを独自に開発し、混合試料の厳密な評価を可能にすることを目指した。ソフトウェアはオープンソースのプログラミング言語 R を用いることで、プログラムの透明性と普及性を確保した。また、計算に使用するパラメータは、多数の実験データに基づいてローカス毎の細かな違いを推定した上でソフトウェアに組み入れた。その結果、開発したソフトウェアは混合試料中に被疑者の DNA が含まれているか否かを高精度で判断することができた。また、試料の DNA が微量な場合でもほぼ問題なく含まれていることを判断できた。さらに、ソフトウェアを米国の標準ガイドラインに基づいて検証し、その質的保証を明示した。よって、作成したソフトウェアは混合試料の解析に有用であることが示された。

以上の研究は、法医学領域での DNA 鑑定の精度の高い評価に貢献するに留まらず、わが国における科学捜査の発展と犯罪抑止につながるものと期待される。

したがって、本論文は博士（医科学）の学位論文として価値あるものと認める。

なお、本学位授与申請者は、平成 30 年 2 月 9 日実施の論文内容とそれに関連した試問を受け、合格と認められたものである。